基于遥感数据的新疆开—孔河流域农业区种植结构提取①

茶明星[1,2,3]， 汪小钦[1,2,3]， 李娅丽[1,2,3]， 邱鹏勋[1,2,3]
（1. 福州大学空间数据挖掘与信息共享教育部重点实验室，福建 福州 350108；
2. 卫星空间信息技术综合应用国家地方联合工程研究中心，福建 福州 350108；
3. 数字中国研究院（福建），福建 福州 350108）

摘　要：农作物种植结构是农作物空间格局的重要组成部分，是区域土地资源和水资源优化配置的基础。以新疆开—孔河流域农业区为研究区域，综合利用作物物候信息和 2016 年的 MODIS NDVI 时序曲线，获得不同作物生长差异明显的关键期，选择关键期的 Landsat 8 OLI 影像，构建主要作物提取知识规则，基于决策树方法开展农作物的分类识别。开—孔河农业区 2016 年主要作物种植面积为 5.07×10^5 hm^2，其中棉花种植面积最大，为 1.97×10^5 hm^2，玉米、小麦次之。博斯腾湖和开都河农业区以辣椒、玉米和小麦为主要作物，种植结构比较零散；孔雀河农业区种植结构比较单一，以棉花和香梨为主要作物。与仅利用时间序列的 MODIS 数据进行作物分类识别的结果对比表明，综合利用 MODIS 和 Landsat 数据的作物识别精度有显著提高，总体分类精度从 62.58% 提高到 88.37%，kappa 系数从 0.53 提高到 0.86。该方法综合利用了 MODIS 数据的时序特征和 Landsat 数据较高的空间分辨率特征，有效地避免了 MODIS 数据空间分辨率不足而导致分类精度较差的情况，又避免了 Landsat 数据时间分辨率不足所引起的时相选择盲目性或数据冗余，在干旱区农业种植结构的提取领域具有一定的应用价值。

关键词：农作物；种植结构；MODIS；Landsat 8；多源数据集成；开都—孔雀河流域

农作物种植结构描述的是一个地区或农业生产单位内农作物的类型及其空间分布[1]，是农作物空间格局的重要组成部分，是区域土地资源和水资源优化配置的基础[2]。准确高效地获取农作物种植结构和主要农作物产量信息不仅对作物种植结构的优化和调整具有重要意义[3]，对于提高农业用水效率以及缓解干旱区水资源紧张也具有重要的意义。

多种农作物通过一定的种植方式组合形成了复杂的农业土地系统，单一的遥感影像难以覆盖所有作物光谱识别的最佳时相，利用时间序列遥感影像表征作物季相及物候变化规律，进而开展作物遥感提取的方法得到了广泛应用[4-5]。高时间分辨率的 MODIS 数据在大面积作物种植信息提取研究中的应用较多[6-10]，实现了小麦、玉米和大豆等农作物的分类识别和种植结构的提取研究。利用时间序列的 MODIS 影像进行作物信息提取能够充分发挥高时间分辨率的优势，但受空间分辨率的限制，只能应用于大面积农作物种植区域，在农作物种植结构相对复杂的区域无法获得理想的提取效果。中等空间分辨率的 Landsat 数据在作物种植信息提取的研究中也有一定的应用[11-12]，实现了水稻等典型作物的信息提取。利用单一中等空间分辨率的 Landsat 影像进行作物分类识别的方法一定程度上降低了存在大量混合像元的现象，提高了作物分类的精度，却难以获得时间序列较为完备和密集的影像数据。

集成多时空分辨率的多源遥感数据能够增强作物在物候和空间等特征上的差异信息，一定程度上克服了作物分类时“同物异谱，异物同谱”现象的发生，有利于较高精度的作物识别。目前，部分学者利用时空融合技术构造高空间和时间分辨率的时序数据集，开展基于多源遥感数据的作物信息提取研究。如陈梦露等[13]在我国北方种植大省河北、黑龙江、新疆典型农区，基于时空数据融合模型，利用 Landsat 和 MODIS 影像构建了高时空分辨率影像数据

---

① **收稿日期：**2019－05－14； **修订日期：**2019－08－20
基金项目：国家重点研发计划课题（2017YFB0504203）；中央引导地方发展专项（2017L3012）
作者简介：茶明星（1997－），女，硕士研究生，研究方向为遥感技术与应用. E-mail：N185527001@ fzu. edu. cn
通讯作者：汪小钦. E-mail：wangxq@ fzu. edu. cn

集,探讨时空数据融合模型在农业信息提取领域的适用性;Hao 等[14]、Laura-Piedelobo 等[15]综合 Landsat 和 Sentinel-2 等数据实现了多源数据的作物信息提取研究。虽然集成多时空分辨率遥感数据在农业信息提取领域具有一定的优势和适用性,但目前尚无一种能够适用于不同数据源和应用目的的时空融合算法出现[16]。已有的时空融合算法都有其各自的优点和局限性,多源遥感影像的几何和辐射特性差异的校正、异质地表混合像元模型的非线性散射特征的处理和地物时相变化模型的建模等过程对不同的数据源和地物的变化会造成不同程度的误差,而且模型比较复杂,不利于方法的推广应用。

本文以新疆开都—孔雀河流域(以下简称开—孔河流域)农业区为研究区,综合利用高时间分辨率的 MODIS 影像数据获得作物差异明显的关键生长时期,选择同时期较高空间分辨率的 Landsat 8 影像,开展作物种植结构的提取研究,以期提供一种简单有效地集成多源数据的样本点选择和作物识别方法,为干旱区农业种植结构信息的提取和应用提供参考。

1 研究区概况与数据源

1.1 研究区概况

研究区位于新疆开—孔河流域(82°56′~88°12′E,40°48′~43°20′N)农业区(图1),主要包括博斯腾湖农业区 1 019 km^2、开都河农业区 2 459 km^2 以及孔雀河农业区 2 789 km^2[17]。开—孔河流域具有明显的大陆性干旱气候特征,总体特点为日照时间长、昼夜温差大、气候波动频繁、降雨稀少、蒸发剧烈。农业区降水主要集中在 4—7 月,最大降水量在 20~40 mm,农作物生长期主要集中在 3—10 月。博斯腾湖、开都河以及孔雀河农业区是该地区重要的特色农产品生产基地,其中博斯腾湖和开都河农业区主要集中在和静县、和硕县、焉耆县和博湖县,主要种植作物有小麦、玉米、棉花、辣椒等;孔雀河农业种植区主要集中在库尔勒市和尉犁县,主要种植作物有棉花和香梨。

1.2 数据源及预处理

研究采用的数据包括遥感数据、主要农作物物候信息和野外考察样本点数据。

遥感数据包括 Landsat 8 OLI 数据和 MODIS 数据(表1),其中 Landsat 8 OLI 数据空间分辨率为 30 m,从地理空间数据云平台(http://www.gscloud.cn/)下载 2016 年 4—9 月的 7 景 Landsat 8 OLI 数据;从 NASA 官网(https://ladsweb.nascom.nasa.gov/search)下载的 2016 年全年 23 期分辨率为 250 m 的 MOD13Q1 以及秋、夏季节的 4 景 500 m 分辨率的 MOD09A1 产品数据。

Landsat 8 OLI 数据预处理主要包括辐射定标、

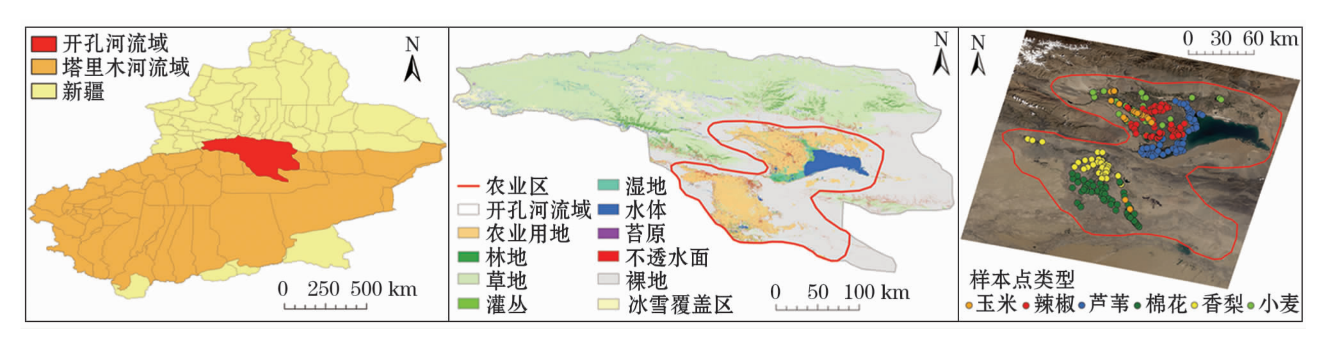


图 1 研究区地理位置及样本点空间分布示意图

Fig. 1 Location of study area and spatial distribution of sample points

表 1 遥感数据列表

Tab. 1 Remote sensing data list

传感器	获取时间/月-日						
Landsat 8	04-14	05-16	06-01	07-27	08-04	09-05	09-21
MOD13Q1	01-01	01-17	02-02	02-18	03-05	03-21	04-06
	04-22	05-08	05-24	06-09	06-25	07-11	07-27
	08-12	08-28	09-13	09-29	10-15	10-31	11-16
	12-02	12-18					
MOD09A1	06-09	07-11	08-12	09-13			

大气校正、影像裁剪等内容。MODIS 数据已经过严格的几何校正、辐射定标、大气校正等预处理,利用 MRT 工具对数据进行投影变换(正弦投影转 UTM_Zone_45N)、格式转换(HDF 转 TIF)、波段提取(NDVI 波段和对应的 QC 波段)、影像裁剪以及 S－G 滤波[18－19]等处理。

从中国气象数据网(https://data.cma.cn/)获得新疆开—孔河流域小麦、玉米、棉花 3 种农作物物候期资料;实地调查获得辣椒、香梨和芦苇物候期资料(表 2)。由于博斯腾湖农业区分布有大量的芦苇,也将其作为农作物分类识别的对象之一。

研究区内农作物样本点数据主要通过实地调查

表 2　开—孔河流域主要农作物物候历

Tab. 2　Main crop phenology of the Kai-kong River Basin

<table>
<tr><td>月</td><td>2</td><td colspan="3">3</td><td colspan="3">4</td><td colspan="3">5</td><td colspan="3">6</td><td colspan="3">7</td><td colspan="3">8</td><td colspan="3">9</td><td colspan="3">10</td></tr>
<tr><td>旬</td><td>下</td><td>上</td><td>中</td><td>下</td><td>上</td><td>中</td><td>下</td><td>上</td><td>中</td><td>下</td><td>上</td><td>中</td><td>下</td><td>上</td><td>中</td><td>下</td><td>上</td><td>中</td><td>下</td><td>上</td><td>中</td><td>下</td><td>上</td><td>中</td><td>下</td></tr>
<tr><td>小麦</td><td colspan="2">播种</td><td colspan="3">出苗</td><td colspan="4">分叶、拔节</td><td colspan="3">拔节、孕穗</td><td colspan="2">乳熟</td><td colspan="2">收割</td><td colspan="9"></td></tr>
<tr><td>玉米</td><td colspan="5"></td><td colspan="2">播种</td><td colspan="3">出苗</td><td colspan="3">拔节</td><td colspan="4">抽穗</td><td colspan="2">乳熟</td><td colspan="2">收获</td><td colspan="4"></td></tr>
<tr><td>棉花</td><td colspan="5"></td><td colspan="2">播种</td><td colspan="5">出苗、长叶</td><td colspan="2">现蕾</td><td colspan="3">花铃</td><td colspan="5">裂铃</td><td colspan="3">棉花采摘</td></tr>
<tr><td>辣椒</td><td colspan="3"></td><td colspan="2">播种</td><td colspan="2">出苗</td><td colspan="3">定植</td><td colspan="4">生长</td><td colspan="3">开花</td><td colspan="4">结果</td><td colspan="4">收获</td></tr>
<tr><td>香梨</td><td colspan="2"></td><td colspan="3">返青</td><td colspan="2">开花</td><td colspan="12">生长</td><td colspan="6">果实成熟</td></tr>
<tr><td>芦苇</td><td colspan="4"></td><td>发芽</td><td colspan="3">展叶</td><td colspan="7">生长</td><td colspan="2">孕穗</td><td colspan="2">抽穗</td><td>开花</td><td colspan="3">成熟</td><td colspan="2">落叶</td></tr>
</table>

并结合 Google Earth 进行采集。2018 年 7 月对研究区内的农作物种植情况采用 GPS 定点及拍照的方式进行了实地调查,采集实地的农作物样本点。此外,结合 Google Earth 上影像的色彩以及纹理等特征增加样本点,并删除 2016 年与 2018 年影像不一致的野外考察样本点。采集的样本点总计 368 个(图 1),其中小麦 67 个、玉米 50 个、棉花 63 个、辣椒 76 个、香梨 42 个、芦苇 70 个,主要用于精度验证。

2　研究方法

为了解决研究区实际样本点无法获取的困难,本文基于 Landsat 8 OLI 和 MODIS NDVI 影像构造作物样本点。进而得到不同作物的 NDVI 年际变化曲线,分析可得作物生长明显差异的关键期,选择关键期的 Landsat 8 影像,通过构造作物识别知识规则,基于决策树分类方法开展作物的识别与分类,并基于单一 MODIS 影像数据进行作物识别与分类的方法进行对比分析,探讨基于多平台遥感数据在农业种植结构提取方面的适用性。

2.1　选取作物样本点

实验中采用的是 2016 年的数据,由于无法获得研究区实际的农作物样本点,本文基于 Landsat 8 和 MODIS 影像,结合不同作物的物候信息构造和筛选作物样本点。

首先,利用专家知识,根据不同作物在同一时期的 NDVI 差异及同种作物在不同时期的 NDVI 变化,选择不同月份的 Landsat 8 影像建立初始样本点。具体方法为(图 2):选择 a 时期影像和 b 时期影像,其中 a、b 时期影像中的每个小矩形代表一部分像元,颜色代表像元的 NDVI 值,NDVI 不同则颜色不同。在影像 a、b 中,可以区分出深绿色和白色为不同的作物,对 a、b 影像进行叠加,则可以区分出红色、橘色、蓝色和紫色 4 种不同的作物。结合农作物物候信息,再对 4 类作物进行区分,确定具体为哪一类作物,直到选择出所有作物样本点。基于该方法,构造的候选样本点总计 394 个,其中小麦 43 个、玉

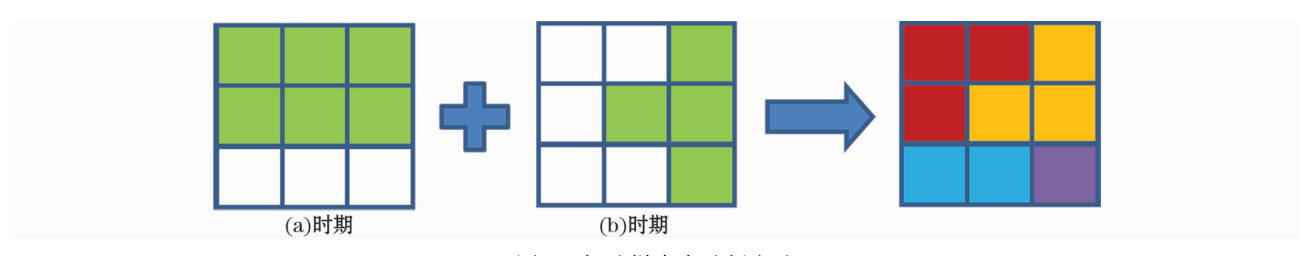


图 2　候选样本点选择方法

Fig. 2　Candidate sample point selection method

米61个、棉花52个、辣椒68个、香梨71个、芦苇99个。完成候选样本点构造后,利用MODIS数据对候选样本点进行筛选。计算394个候选样本点的全年23期MODIS影像的NDVI值进行提取,得到每个候选样本点的NDVI时序变化曲线,剔除每一类候选样本点中NDVI明显异常的曲线和不符合农作物物候规律的曲线,得到符合作物生长NDVI曲线的样本点总计149个,其中小麦16个、玉米29个、棉花15个、辣椒26个、香梨30个、芦苇33个。结合高分辨率的Google Earth影像对筛选得到的样本点进行验证,并与2018年野外考察的样本数据中确认未发生变化的地物样本点进行对比,结果显示最终筛选出的样本点符合实际情况。

2.2 作物生长期NDVI曲线分析

对最终确定的典型作物样本点NDVI曲线求平均得到小麦、玉米、辣椒、棉花、香梨和芦苇的NDVI年际变化曲线(图3)。

由农作物生长期NDVI曲线可知,小麦、香梨和芦苇3种作物的生长期早于辣椒、棉花和玉米,这与作物物候历是吻合的,说明所选择的样本点符合实际情况。小麦、香梨和芦苇3种作物在3月中旬左右开始生长,而辣椒、棉花和玉米3种作物则在4月中旬左右开始种植。据此,我们可以选择3月中旬到4月上旬的关键时期影像将其分为两大类,然后再进行细分。

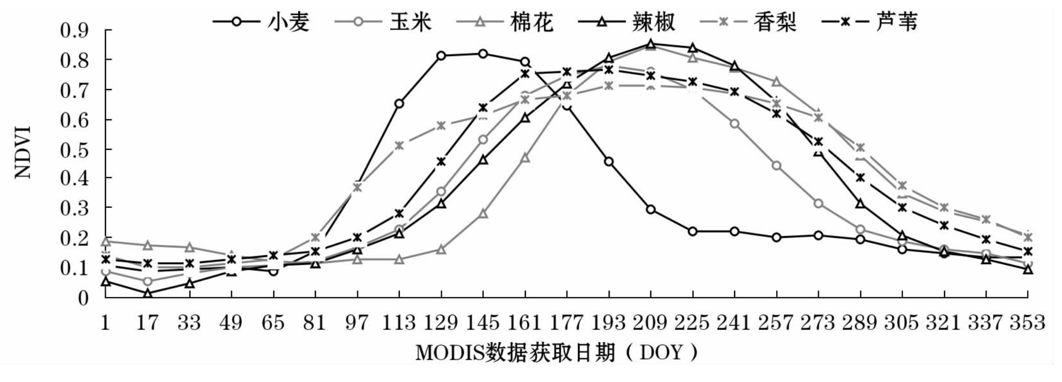


图3 不同作物及易混淆地表覆盖NDVI年内变化曲线

Fig. 3 Annual variation curve of NDVI covered by different crops and confusing surface

2.3 基于决策树分类的作物识别与提取

根据不同作物NDVI年际变化曲线找到差异明显的关键时间节点,并选择该节点所对应的同时期较高空间分辨率的Landsat 8影像通过构建知识规则,可以最大程度地利用较少影像实现NDVI差异明显的作物分类,但对于生长期较为相似的作物,仅基于NDVI无法实现分类。因此,结合作物在不同波段的光谱差异进一步进行分类。同时,为了避免其他信息对农作物提取的干扰,需要预先提取出水体和其他非植被区域。

2.3.1 MNDWI规则 利用MNDWI(Modification of normalized difference water index)[20]对水体进行提取。计算公式如下:

$$MNDWI = \frac{Green - SWIR}{Green + SWIR} \quad (1)$$

式中:SWIR表示短波红外波段的反射率值,Landsat 8 OLI第6波段、Landsat 7 ETM+第5波段的反射率值;Green表示可见光绿色波段的反射率值,Landsat 8 OLI第3波段、Landsat 7 ETM+第2波段的反射率值。为了避免冰雪融化的干扰,选择5月16日(第137 d)、9月5日(第249 d)以及9月21日(第265 d)3个时期的影像分别代表夏初、秋初和秋末来提取水体信息,具体规则为($MNDWI_{5/16}>0.1$) and ($MNDWI_{9/5}>0.1$) and ($MNDWI_{9/21}>0.1$)。

2.3.2 NDVI规则 本文利用NDVI(NDVI-Normalized Difference Vegetation Index)[21]对农作物进行提取。计算公式如下:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2)$$

式中:NIR表示近红外波段的反射率值,Landsat 8 OLI第5波段或Landsat 7 ETM+第4波段的反射率值;Red表示可见光红色波段的反射率值,Landsat 8 OLI第4波段或Landsat 7 ETM+第3波段的反射率值。选择5月16日(第137 d)、7月27日(第209 d)、9月21日(第265 d)三个均有植被覆盖时期的影像来排除裸地,具体规则为:($NDVI_{5/16}>0.1$) or

($NDVI_{9/5}>0.1$)or($NDVI_{9/21}>0.1$)。

2.3.3 *农作物知识规则的建立* 小麦、香梨和芦苇3种作物的生长期早于辣椒、棉花和玉米,5月中旬,小麦处于孕穗期,香梨处于叶子生长期,植被覆盖度很高;而玉米、棉花和辣椒处于出苗期,植被覆盖度很低,NDVI值远低于0.4。选择5月16日(第137 d)的关键期影像将其分为两大类,分类规则为:($NDVI_{5/16}>0.4$)。接下来在每一类中进行细分。

(1) 小麦、香梨、芦苇提取

小麦与香梨和芦苇分离的关键节点在7月下旬,此时小麦已经收割,在7月27日(第209 d)的NDVI值远低于0.3,而此时香梨和芦苇NDVI值在0.5以上,通过设置规则($NDVI_{7/27}<0.5$)将小麦与香梨和芦苇进行分离。由于香梨和芦苇的NDVI生长曲线相似,单一的基于NDVI值无法将两者很好地分离,因此,结合光谱规则进行分离。经过多次实验及反复对比,选择蓝色波段(band blue)和近红外波段(band NIR)的值作为光谱特征参数对香梨和芦苇的空间分布特征进行识别和提取,具体规则为:芦苇($B_{NIR}\leqslant 2\,500$ or $B_{blue}\leqslant 600$)。

(2) 玉米、棉花、辣椒提取

玉米与与其他作物分离的关键节点在9月下旬,此时玉米已经收割,棉花和辣椒仍然处于生长期。在9月21日(第265 d)玉米的NDVI值远低于0.4,而此时棉花和辣椒的NDVI值在0.5以上,通过设置规则($NDVI_{9/21}<0.5$)将玉米与棉花和辣椒进行分离。由于芦苇生长状况参差不齐,利用5月16日影像进行分类时,玉米、棉花和辣椒中也被分进部分芦苇。因此,在分离棉花和辣椒之前需要把芦苇提取出来。选择5月16日蓝光波段(band blue),设置规则($B_{blue}\leqslant 800$)可将芦苇从棉花和辣椒中剔除。9月21日,棉花处于花铃收获期,辣椒处于采摘期,都已停止生长。但由于棉花是白色,辣椒是红色,其光谱特征存在差异。为了准确识别二者的空间分布特征,选择蓝光波段(band blue)和近红外波段(band NIR)的值作为特征参数对辣椒和棉花空间分布特征进行识别和提取,具体规则为:辣椒($B_{blue}\leqslant 550$ or $B_{NIR}\leqslant 550$)。基于分类规则构造的决策树如图4所示。

2.4 分类精度评价方法

混淆矩阵是表示精度评价的一种标准格式,用n行n列的矩阵形式来表示。具体的评价指标有总体精度OA(overall accuracy)、Kappa系数等。

总体分类精度描述的是对每一个随机样本,所分类的结果与参考数据所对应区域的实际类型相一致的概率,计算公式如下:

$$OA=\sum_{i=1}^{n}X_{ii}/N \tag{3}$$

式中:N表示样本点总数;X_{ii}表示被分到正确类别的样本数。

Kappa系数是一种定量评价遥感分类图精度的方法,计算公式如下:

$$Kappa=\frac{N\sum_{i=1}^{n}X_{ii}-\sum_{i=1}^{n}(X_{i+}X_{+i})}{N^2-\sum_{i=1}^{n}(X_{i+}X_{+i})} \tag{4}$$

式中:N表示样本点总数;X_{ii}表示被分到正确类别的样本数;X_{i+}和X_{+i}分别表示第i行和第i列的总样本数量。

3 结果评价与分析

3.1 分类结果

利用图4构建的决策树,对开—孔河农业区的主要农作物进行分类,获得初步分类结果。由于决策树分类是基于像元的作物分类方法,其分类结果不可避免地出现一些像元孤岛等“椒盐”现象,利用分类后处理功能模块对分类结果进行小斑块去除。随后根据当地作物实际种植的经验知识,并借助Goolge Earth影像,对去除小斑块后的结果再进行适当的人机交互处理。由图5a可知,博斯腾湖和开都河农业区作物种植主要为小麦、玉米、辣椒及芦苇。其中芦苇为自然生长,其分布集中在博斯腾湖西部外围区。而小麦、玉米、辣椒分布较零散,没有明显的规律。相比之下,孔雀河农业区主要为棉花和香梨,种植比较集中,只零星分布一些玉米和小麦。开展仅利用时间序列的MODIS影像的农作物种植结构提取,同样基于图5的规则对阈值进行适当调整,经分类后处理得到的农业种植结构提取如图5b。由于MODIS影像的空间分辨率为250 m,混合像元较多,对比图5a和图5b,直观上可以看出分类的效果较差,很大一部分农作物的像元被分到非植被区域。

3.2 分类精度评价

利用野外考察样本点,采用混淆矩阵的方法对上述分类结果进行精度验证(表4)。

综合利用MODIS和Landsat 8影像的分类方法与利用单一MODIS影像的分类方法相比,除香梨提

Landsat 8时序数据

($MNDWI_{5/16} > 0.1$) and($MNDWI_{9/05} > 0.1$)and ($MNDWI_{9/21} > 0.1$) —Y→ 水域; N↓

($NDVI_{5/16} < 0.3$) or($NDVI_{7/27} < 0.3$)or ($NDVI_{9/21} < 0.3$) —Y→ 裸地; N↓

$NDVI_{5/16} > 0.4$

- Y: $NDVI_{7/27} < 0.5$
  - Y: 小麦
  - N: ($B_{NIR7/27} \leqslant 2\,500$) or($B_{Blue7/27} \leqslant 600$)
    - Y: 芦苇
    - N: 香梨
- N: $NDVI_{9/21} > 0.5$
  - Y: (B_{Blue}5/16 ≤ 800)
    - Y: 芦苇
    - N: ($B_{Blue9/21} \leqslant 550$) or($B_{NIR9/21} \leqslant 550$)
      - Y: 辣椒
      - N: 棉花
  - N: $NDVI_{9/05} < 0.5$
    - Y: 玉米
    - N: 其他

图 4　农作物分类决策树

Fig. 4　Decision tree of crop classification

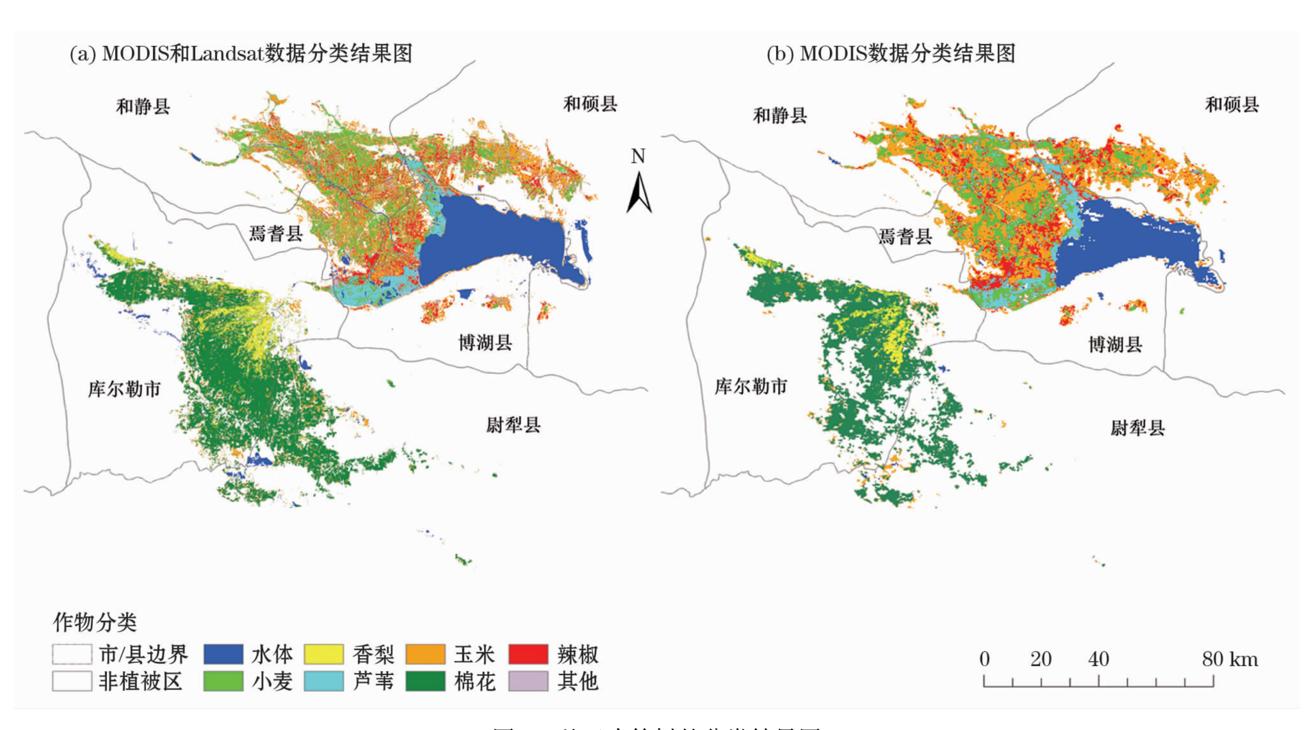


图 5　基于决策树的分类结果图

Fig. 5　Classification result graphs based on decision tree

表 4　分类精度

Tab. 4　Classification accuracy

/%

方法	小麦	玉米	棉花	辣椒	香梨	芦苇	总体精度	Kappa 系数
MODIS + Landsat	78.48	87.50	94.12	85.92	96.43	97.62	88.37	0.86
MODIS	48.48	38.71	64.71	68.42	90.91	84.00	62.58	0.53

取精度略微提高外，其他作物的提取精度都有显著的提高。综合利用 MODIS 和 Landsat 影像进行分类的评价结果，Kappa 系数为 0.86，总体分类精度达到 88.37%。其中芦苇、香梨、棉花分类精度较高，分别为 97.62%、96.43% 及 94.12%。而小麦、辣椒与玉米分类精度相对降低。主要是因为在 30 m 分辨率的 Landsat 8 影像上，3 种作物的种植结构零散且复杂多样，造成混合像元的出现。此外，小麦分类精度较辣椒和玉米更低的原因，是小麦收割以后的地块上会种植蔬菜等其他作物，导致小麦收割后的 NDVI 值仍然处于有植被状态，从而产生误分。而基于单一 MODIS 影像进行分类的评价结果 Kappa 系数为 0.53，总体分类精度只有 62.58%。其中芦苇和香梨分类精度相对较高，分别为 84% 和 90.91%，而其他作物的分类精度则较低，其中玉米的分类精度最低，仅为 38.71%，主要是大量的玉米像元被误分为小麦。由于 MODIS 影像的空间分辨率较低，在 250 m 分辨率的影像上存在着大量的混合像元，这是导致作物分类精度低最主要的原因。

3.3　主要作物种植面积分析

在综合利用 MODIS 和 Landsat 影像进行分类的基础上，进一步对各县市的主要农作物种植面积信息进行统计分析（图 6）。

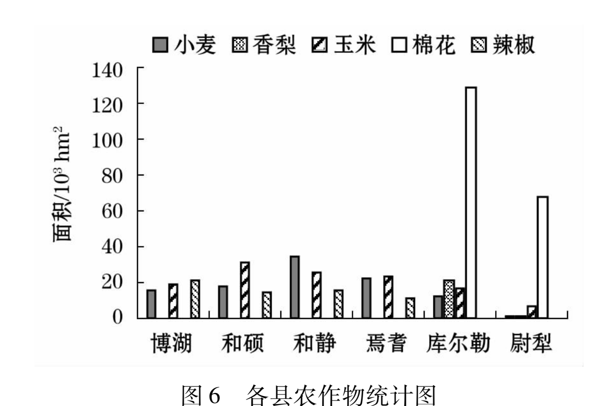


图 6　各县农作物统计图

Fig. 6　Crop statistics of each county

开—孔河农业区主要作物种植总面积为 5.07×10^5 hm²。其中棉花种植面积最大，为 1.97×10^5 hm²。其次是玉米和小麦，种植面积分别为 1.23×10^5 hm² 和 1.03×10^5 hm²，辣椒种植面积为 6.17×10^4 hm²，香梨种植面积最小为 2.28×10^4 hm²。由图 6 可知，孔雀河农业区以棉花种植为主，辅以香梨、玉米、小麦和辣椒等作物，约占整个开—孔河农业区主要作物种植面积的 43.39%。库尔勒市和尉犁县的棉花种植面积均占绝对优势，香梨主要分布在库尔勒市。博斯腾湖和开都河农业区小麦、玉米和辣椒的种植面积较为均衡，约占整个开—孔河农业区作物种植面积的 56.61%。

和硕、和静和焉耆地区作物种植结构比较分散，地块较为破碎，30 m 分辨率的 Landsat 影像上仍然存在一定的混合像元，农田地块之间的防护林像元会被误分为农作物，使得作物种植结构的统计结果出现一定的偏差。小麦收割以后的地块上会种植蔬菜等其他作物，导致小麦收割后的 NDVI 值仍然处于有植被状态，从而产生误分，使得和硕、焉耆等县市小麦的统计结果低于实际种植信息。在今后的工作中，可以考虑结合高分辨率的遥感数据开展作物种植信息的提取。

4　结 论

（1）提出了综合利用作物物候期、多时期 Landsat 8 及时序 MODIS NDVI 影像的信息样本点选取方法，为不易获得样本点年份和区域农作物的提取提供了较好的解决方案。

（2）利用主要作物 MODIS NDVI 时序曲线，找到不同作物差异明显的关键时期，进而选择关键时期较高空间分辨率的 Landsat 8 数据，综合考虑作物的 NDVI 差异以及光谱差异规则，基于决策树的分类方法对作物种植结构较为复杂区域的作物识别能够取得较为理想的分类结果。该方法可以充分发挥两种数据的优势，弥补 MODIS 数据空间分辨率，及 Landsat 8 数据难易对多种作物进行区分的不足，简单高效地实现对作物种植信息的提取。

（3）开—孔河流域农业区主要作物种植面积为 5.07×10^5 hm²。其中棉花种植面积最大为 1.97×10^5 hm²、玉米种植面积为 1.23×10^5 hm²、小麦种植

面积为 1.03 × 10^5 hm^2、辣椒种植面积为 6.17 × 10^4 hm^2、香梨种植面积为 2.28 × 10^4 hm^2。博斯腾湖和开都河农业区约占整个开孔河农业区作物种植面积的 56.61%,以小麦、玉米和辣椒为主,作物种植较为零散。孔雀河农业区约占整个开孔河农业区作物种植面积的 43.39%,作物种植比较单一,以棉花和香梨为主。

参考文献(References):

[1] 唐华俊,吴文斌,杨鹏,等. 农作物空间格局遥感监测研究进展[J]. 中国农业科学,2010,43(14):2879 - 2888. [Tang Huajun, Wu Wenbin, Yang Peng, et al. Recent progresses in monitoring crop spatial patterns by using remote sensing technologies[J]. Scientia Agricultura Sinica,2010,43(14):2879 - 2888.]

[2] 蔡超,夏建新,任华堂. 基于蓝水资源的新疆农业种植结构调整分析[J]. 农业现代化研究,2015,36(2):265 - 269. [Cai Chao, Xia Jianxin, Ren Huatang. Blue water oriented optimization of plantation industry in Xinjiang[J]. Research of Agricultural Modernization,2015,36(2):265 - 269.]

[3] 胡琼,吴文斌,宋茜,等. 农作物种植结构遥感提取研究进展[J]. 中国农业科学,2015,48(10):1900 - 1914. [Hu Qiong, Wu Wenbin, Song Xi, et al. Recent progresses in research of crop patterns mapping by using remote sensing[J]. Scientia Agricultura Sinica,2015,48(10):1900 - 1914.]

[4] Elodie Vintrou, Desbrosse Annie, Bégué Agnès, et al. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products[J]. International Journal of Applied Earth Observation & Geoinformation,2012,14(1):83 - 93.

[5] 胡琼. 基于时序 MODIS 影像的农作物遥感识别方法研究[D]. 北京:中国农业科学院,2018. [Hu Qiong. Research on Crop Remote Sensing Recognition Method Based on Time-series MODIS Image[D]. Beijing: Chinese Academy of Agricultural Sciences, 2018.]

[6] Lorenzo Busetto, Zwart Sander-J, Boschetti Mirco. Analysing spatial-temporal changes in rice cultivation practices in the Senegal River Valley using MODIS time-series and the PhenoRice algorithm[J]. International Journal of Applied Earth Observation and Geoinformation,2019,75:15 - 28.

[7] Qiu B, Lu D, Tang Z, et al. Automatic and adaptive paddy rice mapping using Landsat images: Case study in Songnen Plain in Northeast China[J]. Science of the Total Environment,2017,598: 581 - 592.

[8] 宫诏健,刘利民,陈杰,等. 基于 MODIS NDVI 数据的辽宁省春玉米物候期提取研究[J]. 沈阳农业大学学报,2018,49(3): 257 - 265. [Gong Zhaojian, Liu Limin, Chen Jie, et al. Phenophase extraction of spring maize in Liaoning Province based on MODIS NDVI data[J]. Journal of Shenyang Agricultural University,2018, 49(3):257 - 265.]

[9] 周思,何祺胜,刘宝柱. 基于 MODIS 的黑龙江省农作物种植结构提取研究[J]. 地理空间信息,2018,16(1):79 - 82. [Zhou Si, He Qisheng, Liu Baozhu. Extraction of crops planting structure based on MODIS in Heilongjiang Province[J]. Geospatial Information,2018,16(1):79 - 82.]

[10] 郭昱杉,刘庆生,刘高焕,等. 基于 MODIS 时序 NDVI 主要农作物种植信息提取研究[J]. 自然资源学报,2017,32(10):1808 - 1818. [Guo Yushan, Liu Qingsheng, Liu Gaohuan, et al. Extraction of main crops in Yellow River Delta based on MODIS NDVI time series[J]. Journal of Natural Resources,2017,32(10):1808 - 1818.]

[11] 刘吉凯,钟仕全,梁文海. 基于多时相 Landsat8 OLI 影像的作物种植结构提取[J]. 遥感技术与应用,2015,30(4):775 - 783. [Liu Jikai, Zhong Shiquan, Liang Wenhai. Extraction of crop planting structure based on multi-temporal Landsat8 OLI image[J]. Remote Sensing Technology and Application, 2015, 30(4): 775 - 783.]

[12] 阳雄伟,陈春花. 基于融合时间特征 Landsat 影像的水稻面积提取[J]. 地理空间信息,2018,16(6):87 - 89. [Yang Xiongwei, Chen Chunhua. Paddy fields area extraction based on landsat data with integrating temporal features[J]. Geospatial Information, 2018,16(6):87 - 89.]

[13] 陈梦露,李存军,官云兰,等. 基于 ESTARFM 模型的区域农田高时空分辨率影像产生与应用[J]. 作物学报,2019,45(7): 1099 - 1110. [Chen Menglu, Li Cunjun, Guan Yunlan, et al. Generation and application of high temporal and spatial resolution images of regional farmland based on ESTARFM model[J]. Acta Agronomica Sinica,2019,45(7):1099 - 1110.]

[14] Hao P, Wang L, Niu Z. The potential of time series merged from Landsat - 5 TM and HJ - 1 CCD for crop classification: a case study for bole and Manas Counties in Xinjiang, China[J]. Remote Sensing,2014,6(8):7610 - 7631.

[15] Laura-Piedelobo A, David Hernández-López, Ballesteros Rocío. Scalable pixel-based crop classification combining Sentinel - 2 and Landsat - 8 data time series: Case study of the Duero river basin [J]. Agricultural Systems,2019,171:36 - 50.

[16] 黄波,赵涌泉. 多源卫星遥感影像时空融合研究的现状及展望[J]. 测绘学报,2017,46(10):1492 - 1499. [Huang Bo, Zhao Yongquan. Research status and prospect of spatiotemporal fusion of multi-source satellite remote sensing imagery[J]. Acta Geodaetica et Cartographica Sinica,2017,46(10):1492 - 1499.]

[17] 李卫红,陈忠升,李宝富. 新疆开都—孔雀河流域绿洲需水量与稳定性分析[J]. 冰川冻土,2012,34(6):1470 - 1477. [Li Weihong, Chen Zhongsheng, Li Baofu. Analysis of water demand and stability for oasis in Kaidu-Kongque River Basin, southern Xinjiang [J]. Journal of Glaciology and Geocryology, 2012, 34(6): 1470 - 1477.]

[18] Vivótruyols G, Schoenmakers P-J. Automatic selection of optimal Savitzky-Golay smoothing [J]. Analytical Chemistry, 2006, 78 (13):4598.

[19] 李爱农,宋孟强,马利群. MODIS 植被指数时间序列 Savitzky-Golay 滤波算法重构[J]. 遥感学报,2010,14(4):725 - 741. [Li Ainong, Song Mengqiang, Ma Liqun. Reconstruction of NDVI time-series datasets of MODIS based on Savitzky-Golay filter [J].

Journal of Remote Sensing,2010,14(4):725 - 741.]
[20] 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J]. 遥感学报,2005,9(5):589 - 595.[Xu Hanqiu. A study on information extracting of water body with the modified normalized difference water index (MNDWI)[J]. Journal of Remote Sensing,2005,9(5):589 - 595.]
[21] 符思涛,周云. 基于遥感影像的归一化植被指数算法研究[J]. 江西测绘,2010(3):17 - 33.[Fu Sitao,Zhou Yun. Research on normalized vegetation index algorithm based on remote sensing image[J]. Jiangxi Cehui,2010(3):17 - 33.]

Crop planting structure extraction based on remote sensing data in Kai-Kong River Basin, Xinjiang

CHA Ming-xing[1,2,3], WANG Xiao-qin[1,2,3], LI Ya-li[1,2,3], QIU Peng-xun[1,2,3]

(1. *Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education, Fuzhou University, Fuzhou* 350108, *Fujian, China*;
2. *National & Local Joint Engineering Research Center of Satellite Geospatial Information Technology, Fuzhou* 350108, *Fujian, China*;
3. *Academy of Digital China (Fujian), Fuzhou* 350108, *Fujian, China*)

Abstract: Crop planting structure plays an important role in crop spatial pattern, and it is the foundation for the optimal allocation of regional land and water resources. In this paper, taking the Kai-Kong River Basin agricultural area in Xinjiang as the study area, a method for crop classification that comprehensively utilizes high temporal resolution MODIS images and high spatial resolution Landsat images was proposed. Due to large scope of study area, it is difficult to obtain crop sample points uniformly throughout whole research area with traffic and time constraints. MODIS and Landsat data for 2016, combined with crop phenology data, were used to construct experimental sample points; thereby providing a better solution for crop extraction in this area wherein sample access is difficult. NDVI time-series curves for different crops were constructed based on experimental sample points. Based on the NDVI time-series curves, the critical period of crops during growing season were obtained. For these key periods, Landsat 8 OLI images were selected. Next, extraction knowledge rules for the main crops were constructed, and identification and classification of crops were performed based on decision tree. In 2016, main crop planting area was 5.07×10^5 hm^2 of the Kai-Kong River Basin agricultural region, with largest planting area found for cotton (1.97×10^5 hm^2), followed by those for corn and wheat. Bosten Lake and Kaidu River agricultural area was dominated by pepper, corn, and wheat, and the planting structure was relatively scattered. Planting structure of the Peacock River agricultural area was relatively simple, with cotton and pear as main crops. A comparative experiment based on time-series MODIS images for crop recognition and classification was also conducted. Results were verified and compared with sample points of field survey. The accuracy of crop classification using MODIS and Landsat data was obviously improved as compared with the accuracy of crop classification using only time-series MODIS data. Overall classification accuracy increased from 62.58% to 88.37%, and kappa coefficient increased from 0.53 to 0.86. The use of high temporal resolution MODIS data and high spatial resolution Landsat data can improve the accuracy of crop extraction to a certain extent; this avoided (1) poor classification accuracy caused by the insufficient spatial resolution of MODIS data and (2) phase selection blindness or data redundancy caused by the insufficient resolution of Landsat data. Therefore, this approach has high potential application value in the extraction of crop planting structure in arid areas.

Key words: crop; planting structure; MODIS; Landsat 8 OLI; Multi-source data integration; Kaidu-Kongque River Basin